
Special issue

Miguel Paisana

<https://orcid.org/0000-0002-6316-8092>

Miguel.paisana@obercom.pt

Iscte - University Institute of Lisbon,
CIES-Iscte

Paulo Couraceiro

<https://orcid.org/0000-0002-9063-2081>

paulo.couraceiro@obercom.pt

University of Minho, CECS-UMinho

António Vasconcelos

<https://orcid.org/0009-0001-9935-0953>

antonio.vasconcelos@obercom.pt

Iscte - University Institute of Lisbon,
CIES-Iscte

Gustavo Cardoso

<https://orcid.org/0000-0003-2067-4718>

Gustavo.cardoso@obercom.pt

Iscte - University Institute of Lisbon,
CIES-Iscte

Submitted

November 25, 2025

Approved

May 22, 2026

© 2026

Communication & Society

ISSN 0214-0039

E ISSN 2386-7876

www.communication-society.com

2026 – Vol. 39 (1)

pp. 560-579

How to cite this article:

Paisana, M., Couraceiro, P., Vasconcelos, A., & Cardoso, G. (2026). 'Grok, is this true?': Patterns of AI chatbot use in online political discourse on X during the 2025 Portuguese elections, *Communication & Society*, 39(1), 560-579

<https://doi.org/10.15581/003.39.1.036>

'Grok, is this true?': Patterns of AI chatbot Use in Online Political Discourse on X during the 2025 Portuguese Elections

Abstract

Citizens increasingly reference AI tools in online political discussions, yet the communicative functions of these interactions remain underexplored. This study examines how Grok AI was mobilised by users on X ahead of the Portuguese legislative elections of May 2025. Using a grounded theory approach, we analysed 525 replies to posts of major political parties and leaders that mention (@) Grok AI. Most mentions were negatively framed, with only 1% showing evidence of positive tone. Beyond knowledge-building or discussions about electoral performance (eg. polling), users resort to Grok for humour and, more prominently, to moralise political actors and stereotype political identities. However, nearly half of the mentions exhibited an undetermined tone or intention, reflecting the inherent ambiguity and complexity of citizen discourse when evoking Grok as participant. Our findings point towards four distinct use models: (i) opinion validation, (ii) argumentation by proxy, (iii) constructive autonomisation and (iv) discursive and memetic framing.

This study advances research on AI mediated political communication by concluding that citizens primarily appropriate Grok AI as a rhetorical resource to augment their visibility in online political debate, rather than for fact-checking purposes.

Keywords

Artificial Intelligence; Grok AI; X; Twitter; Elections; Social media; Political discussion; Political leaders and parties; Citizen engagement; Uses and affordances.

1. Introduction

The centrality of social media platforms for political discussion, campaigning and influence is particularly salient in the case of X, an aspect that was central to the idea of Twitter and made crucial with the rebranding of the network. The Portuguese "X-sphere" is particularly small, with only 10% of people saying they use X (Cardoso et al., 2025b). However, the network attracts people who are more interested in politics, as well as

political agents who display varying levels of interaction (Ferro-Santos et al., 2024; Gonçalves, 2023)

Like other spheres of human activity and interaction, social media platforms have become experimental fields for the impact of AI in day-to-day life. There is a considerable amount of chatbot-like AI solutions competing for user attention and in this context GrokAI stands out due to its organic integration with X. The ability of the tool to accomplish elaborate tasks through complex reasoning is widely recognised (Jedrzejczak et al., 2025) as is its memetic use as a discursive means to express identity, irony or critique, beyond a mere utility, commonly found within the inner symbology of platforms (Bennet & Segerberg, 2012). While its capacity for neutral and impartial fact-checking has been widely debated, the “Grok is this true?” question has acquired significant cultural meaning beyond mere affordance. It became a “shareable cultural template” (Galip, 2024) within the X community but also recognised beyond it.

The “@Grok” mention functionality in X was made available in Portugal on March 7th, 2025, about two months before the Legislative Election of May 18th, the fourth election since 2019. Portugal is undergoing significant political changes, with the contraction of political cycles, the rise of right-wing populism and the reconfiguration of power dynamics. The 2024 and 2025 general elections put an end to 50 years of bipartisanship between the main center-right and center-left parties (PSD – Partido Social Democrata and PS – Partido Socialista) (Cardoso et al, 2025a). The rise of right-wing populist party Chega and its leader André Ventura has had a profound impact on political debating on social media.

The change in political debate is particularly visible on X, where the memetic tone of the discourse is now entangled with populist and polarising ideas, and is visible in the way that Portuguese users interact with politicians and parties, often in provocative and convoluted terms. Since March 7th, GrokAI became part of that dynamic, allowing users to mobilize a third party into the discussion, for a wide variety of purposes.

In this paper, we explore the early uses of Grok AI in political discourse, focusing on how it was used by citizens to interact with posts by political parties and leaders on X, ahead of the 2025 Legislative elections in Portugal. From this process of observation and analysis, we propose a taxonomy and typology that can help future research to characterize how AI chatbots are integrated into day-to-day political discourse on social media.

Based on a dataset of 525 replies to posts by political parties / leaders that mention “@Grok”, between March 7th and May 18th 2025 we established the following research questions:

RQ1. How was Grok used to interact with posts made by political parties and leaders on X ahead of the 2025 legislative elections in Portugal?

RQ2. To what extent does the usage reveal autonomous fact-checking practices?

2. Literature review

2.1. Social media, disinformation and fact-checking in the electoral context

Social media platforms have become central arenas for political communication, campaigning, and public deliberation. Characteristics, such as instant dissemination and virality reshape political information flows and the relationship between political elites, journalists, and citizens (Chadwick, 2017; Vaccari & Valeriani, 2021). These dynamics open new spaces for participation and visibility but also create vulnerabilities. Platforms are not neutral infrastructures as algorithmic curation and engagement-driven ranking privilege

emotionally charged and polarising content (Tufekci, 2015, van Dijck et al., 2018), becoming fertile ground for misinformation and disinformation (Recuero, 2025).

Disinformation campaigns during elections can influence perceptions on candidates, undermine trust in institutions, and deepen polarisation (Allcott & Gentzkow, 2017; Benkler et al., 2018; Tucker et al., 2018). Electoral disinformation is particularly damaging because it often targets not only political or policy issues but also the legitimacy of electoral processes themselves, such as claims of fraud or miscounts. Such narratives can erode confidence in democratic outcomes and discourage voter participation (Ognyanova et al., 2020, Nisbet et al., 2021, Lewandowski, 2023) and have led to the promotion of fact-checking interventions and counter-messaging strategies (OECD, 2024, UN, 2022).

In Europe, these challenges have been addressed through a combination of regulatory frameworks, coordinated monitoring, and collaborative fact-checking initiatives. The European Union Code of Practice on Disinformation (2018, strengthened in 2022) established a self-regulatory mechanism whereby major platforms commit to transparency reporting, advertising restrictions, and cooperation with fact-checkers (European Commission, 2025).

Mis/disinformation has been a salient feature of European elections, from the 2019 European Parliament elections to recent national contests in countries such as France, Spain, and Italy (Paisana et al., 2025). The European Digital Media Observatory (EDMO), launched in 2020, strengthened the role of independent hubs across member states in detecting and countering disinformation during electoral periods, including through cross-border fact-checking networks. Independent hubs such as IBERIFIER, covering Spain and Portugal, BECID (Estonia, Latvia and Lithuania) or NORDIS (Norway, Sweden, Denmark and Finland) have been instrumental in creating better understanding of the transnational scale of disinformation in Europe, highlighting actors, narratives and discursive structures.

In Portugal, electoral disinformation has become a prevalent topic amidst electoral campaigns at least since the 2019 legislative elections. During the 2025 Legislative Election, the focus of disinformation narratives shifted to immigration rather than corruption, traditionally one of the most prevalent disinformative topics (Cardoso et al., 2025a). With increasing political polarisation and short political cycles of government, there is an increased focus on sensible and divisive political issues that require added attention to disinformation and frequent mobilisation of fact-checking skills (Paisana et al., 2025).

Fact-checking has become more common and is increasingly viewed as a necessary journalistic and civil practice, particularly regarding the political sphere (Graves, 2016). It has gained relevance as a vital part of the journalistic practice, requiring the expansion of human and technological infrastructure around it (Juneja & Mitra, 2022, Suomalainen et al., 2025). However, it also increased the level of political instrumentalisation of journalistic fact-checking (Li et al., 2022) and has nudged users to become ad-hoc fact-checkers themselves in online environments (Graves, 2016). In social media environments, the ability to mobilise fact-checking skills and narratives becomes vital to the performance and status of users as debaters (Papacharissi, 2014; Benkler et al., 2018; Farkas & Schou, 2023).

According to Digital News Report data (Cardoso et al., 2025b) the Portuguese adhere to fact-checking in a very similar way to individuals in other countries. To verify information about 1 out of 3 rely on news brands they trust, official sources (e. g., Government website) or search engines. Only 1 out of 10 people resort to AI chatbots although this proportion doubles among 18–24’s and is even higher among younger males. The consolidation of fact-checking practices around news sources and journalists is closely tied to trust in those

sources an aspect that helps justify the high reliance on news brands for this purpose, as Portugal consecutively ranks among the countries where people trust news the most (Cardoso et al., 2025b, Newman et al., 2025).

While evidence and fact-based arguments are positive, weak fact-checking structures or procedures, sustained by algorithms or artificial intelligence may spawn negative side-effects and biases such as the “nobody-fools-me-perception” (Martínez-Costa et al., 2022, Chen, 2025).

This discussion falls within the scope of the debate on the threats and opportunities posed by AI, at a societal level (Qian et al., 2024, Roa, 2024) and in the political sphere (Islam et al., 2024, Chehoudi, 2025). Recent research identifies three uses of AI in political campaigns: campaign operations, voter outreach, and deception (Jungherr et al., 2026). As the authors note, although AI is generally viewed negatively when used in campaigns, this perception is especially strong when it is used deceptively. Nevertheless, even when political actors use AI in this manner, their popularity typically remains unchanged.

The proneness of AI models for sycophancy is likely to confirm the biases and world views of prompters, rather than confront them with opposing ideas or arguments (Sharma, 2023, RRV, 2024) and may also increase the risk of misinformation due to the compliance with illogical or misleading user requests (Chen, 2025). Therefore, the user experience and role in the debate is subject not only to the willful actions in extracting legitimate information, but also to the risk of increased confirmation bias in a hyper-personalised echochamber. Despite the risks, AI chatbots are legitimate tools for fact-checking information, including political content, when used in appropriate circumstances, although some variation may occur depending on factors such as the prompt and language used (Kuznetsova et al., 2025).

2.2. Platforming AI: affordances and AI integration on X

Interest in the microblogging platform X, formerly known as Twitter, has increased in recent years due to changes in its ownership and governance. Discussions on X, which are public by default (Bruns, 2012), typically occur using the “@” sign, allowing a post to be directed at another user, regardless of any pre-existing connection or prior permission (Murphy, 2018).

X has low penetration in Portugal compared to other social platforms (around 10%). However, it attracts a slightly higher percentage of users in the 18–24 age group, particularly those interested in politics. It is also reported that attention to politicians is higher on X than on other platforms (Cardoso et al., 2025b) and during electoral periods in the country, discussions about politics and elections are considerably more frequent on X (Cardoso et al., 2025a).

Studies on political communication highlight that, despite the platform’s potential for interaction with citizens, politicians generally do not interact directly with their constituents (Waisbord & Amado, 2017), often adopting a broader strategy of information and news dissemination (López-García et al., 2016). This tendency is also evident in how political actors in Portugal use the X platform, even though there are important differences between national political parties in the levels of interaction displayed on X (Gonçalves, 2023; Ferro-Santos et al., 2024).

In 2022, X (then Twitter), underwent its most significant change since its public launch when it was formally acquired by Elon Musk. Studies indicate a subsequent increase in hate speech (Benton et al., 2022) and greater popularity, in terms of engagement, of contentious actors (Barrie, 2022). Following the acquisition of the company, there was a shift from

predominantly centralised or third-party content moderation mechanisms to a system primarily based on user-generated “Community Notes” (Jude & Matamoros-Fernández, 2025).

Community Notes “(...) allows people to identify information in Tweets they believe is misleading and write notes that provide informative context” (X Corp, 2021). Whether a note is published is determined by a “bridging algorithm” that prioritises notes with some level of consensus or agreement from users on different sides of the political spectrum (Wojcik et al., 2022).

Research on Community Notes has shown that while users tend to agree with community-based fact checks (Drolsbach & Pröllochs, 2023), these have no meaningful effect on reducing engagement with misinformation because are too slow to deploy and don’t impact engagement in the first stages of misinformation spread (Chuai et al., 2024). Additionally, it is estimated that around 90% of notes are never published, as they do not achieve the necessary consensus from users (Fortune, 2025), which complicates fact-checking practices on more polarising issues.

In November 2023, Grok was integrated into X. It was developed by xAI, founded by Elon Musk in 2023, and was released in the EU and Portugal, in May 2024. Grok is described as an “AI assistant who helps complete tasks, like answering questions, solving problems, and brainstorming” (X Corp, n.d.). Its most touted features are its ability to analyse data directly from X in real time and its so-called “rebellious streak” (X Corp, n.d.).

As noted by Samet (2025), “this direct, low-friction access to an AI assistant that can pull real-time data from the very platform where misinformation often circulates positions Grok as a readily available first-line tool for verification for X users” (p. 1731). However, criticism has also emerged that this feature might cause Grok to amplify false or misleading content already on the platform, reflecting X’s biases more than other LLMs trained on less specific data (Al Jazeera, 2025).

Grok’s “rebellious streak” means it adheres less to the political correctness surrounding other LLMs. Grok tends to provide more unfiltered responses and addresses topics and discussions that other LLMs avoid due to user safeguards. Consequently, this also means that Grok’s “hallucinations”, which are common among most LLMs, also tend to be more extreme (Christopher & Pepe, 2025; Isaac, 2025).

In recent research, Christopher and Pepe (2025) conducted an exploratory analysis of how users interact with Grok, focusing on Gavin Newsom’s posts about the LA protests and Elon Musk’s allegations regarding Donald Trump’s involvement with Jeffrey Epstein. They found that while Grok was primarily used to understand the context of the posts and to conduct fact checks, about 20% of posts concerning Musk’s tweet consisted of “trolling”. This was also common in Gavin Newsom’s posts, with users making requests such as “narrate the LA protests in pirate-speak”. Another key point is that about 10% of users “shared their opinion”, questioning Musk’s motives and credibility, and wanted Grok’s judgement or speculation on possible futures of Musk-Trump fallout” (Christopher & Pepe, 2025).

Overall, while Grok has yielded positive results in terms of accuracy, for example in clinical and educational settings (Jedrzejczak et al., 2025), questions remain about its efficiency and fact-checking potential, and how users appropriate the tool, which diverges from other LLMs due to its controversial approach.

3. Methods

With the objective of exploring the early uses of Grok AI in political discourse ahead of the Portuguese 2025 Legislative elections, we collected reply posts from X that used the “Ask Grok” function while tagging political parties or their political leaders.

Covered politicians include André Ventura (Chega – far-right populist party), Inês Sousa Real (PAN – political center niche party, connected to animal rights and environmental causes), Luís Montenegro (PSD – Centre right legacy party, with ample governative experience), Mariana Mortágua (BE, far-left party connected to identitarian causes), Pedro Nuno Santos (PS – Centre-left legacy party, with ample governative experience), Rui Rocha (IL – Liberal right-wing political party)) and Rui Tavares (Livre – ecological left-wing political party) and covered parties were Bloco de Esquerda (BE), Iniciativa Liberal (IL), Livre, Chega, Partido Comunista and Partido Socialista (PS). This selection comprises all parties with representation in Parliament as elected by voters having official presence on X and that were targeted with posts or replies with “@Grok”. Notable exclusions are Paulo Raimundo, leader of the Portuguese Communist Party, who is not on X, and party accounts for PSD and PAN, which were not mentioned with “@Grok”.

525 reply posts were collected using ZenScraper (oDay3xploit, 2025), which were posted from March 7th, 2025, up until May 18th, the day of the 2025 General Election in Portugal. March 7th is the date X enables the “Ask Grok” feature in replies across several European countries, including Portugal (Mehta, 2025).

These 525 posts did not simply mention Grok, they actively evoked its participation in the conversation, either by using the “Ask Grok” function or by prompting Grok in replies directed at political parties or politicians. As such, the dataset reflects both the initial phase of public experimentation with the feature and the broader electoral discourse leading up to the elections.

The development of a taxonomy (as seen in Table 1) was guided by a grounded theory approach (Glaser & Strauss, 1967, Charmaz, 2006) in an iterative process combining individual research notes, shared reflections and analytical work with the data. The posts were anonymized and imported into a spreadsheet and the three coders coded an initial sample of 105 posts (20% of the overall sample) following the flexible framework suggested by Braun and Clarke (2006): familiarisation with data, generation of initial codes, search for common typologies or traits and subsequent review and naming of the categories.

In the first stage, a free coding procedure was employed with coders independently attaching descriptive labels to each post without being constrained by a predefined coding scheme. This exploratory phase resulted in an initial set of 22 categories, which were subsequently grouped along three provisional analytical dimensions: tone, intent, and content. Through comparative discussions, several categories were consolidated to reduce redundancy and to identify points of agreement and disagreement among coders.

As the analysis progressed, it became clear that content cues were systematically mobilised by coders to infer communicative intent, suggesting a high degree of overlap between these two dimensions. Content was therefore dropped as an autonomous category, and the framework was simplified to focus exclusively on tone and manifest intent. In a further step aimed at ensuring mutual exclusivity and coder reliability, the team agreed to anchor the coding process in tone as the primary entry point, from which manifest intent was then inferred. This methodological decision explains why all intent categories are defined in relation to tone.

Based on this collaborative process we constructed a taxonomy to classify interactions with Grok on X for engagement with political actors, following a bidimensional approach focused on a) *Tone* of the interaction and b) *Manifest intent* of the user.

Table 1. Codebook for characterizing Grok usage in replies to politics related posts on X.

Dimension	Category	Categorial variations	Definition	Examples
Tone	Neutral	Neutral / Curious	Objective language, seeking information and context	@Grok is this true? @Grok is this statement true?
	Positive	Supportive	Affirmation, support or validation of a preferred political figure or idea.	@Grok the performance of [Party leader X] in the debate was penalised by a biased moderator. What is the situation really like regarding topic [Y]
	Negative	Aggressive, insulting or hostile tone. Grok is used to attack or delegitimize a specific politician/party.		@Grok explain to [Party leader X] why he / she and the party are the biggest fake-news spreader and the most corrupt people in the country @Grok is there someone in the country more dishonest than [Party leader X]? @Grok explain [Party leader X] why policy [Y] doesn't work giving examples from country [Z] @Grok explain [Party leader X] how what he / she posted is incoherent with what he/she said
	Undetermined	----	Comment-like interaction with the original post without clearly stating a positive or negative judgement.	@grok can you explain to me why after all the revelations surrounding [Party Leader X], he can still get Y% of the votes? @grok [Party Leader X] writes for the [Y] newspaper. Which group owns this newspaper? Are there any news reports about the administrators' paychecks?
	Knowledge-building	----	Tone tends to be neutral, seeking information or verifying existing information or knowledge.	@Grok analyse in detail and impartially what is said in this video and confirm if it's true @Grok What is the approach of this book? What are the general ideas of the book?
Manifest Intent	Stereotyping	----	Tone is always negative, with ideological characterization and stereotyping, relying on comparative negative statements mostly focusing on past regimes and political events	@Grok in what countries were Karl Marx's ideas successful and societies prospered? @Grok are all fascist regimes based on marxist dogmas and ideologies?
	Moralising	----	Tone is always negative, addressing personal or political misconduct, frequently based on legal issues, corruption, ethics, or morality.	@Grok who is the most lying politician in Portugal? Just say his / her name. @Grok how many lies does [Party leader X] tell per day?
	Electoral performance	----	Tone tends to be neutral, with references to polls, vote %'s, ratings and prediction markets.	@Grok is this poll trustworthy? @Grok What polling company is this? Can it be trusted? Where can I get further information?
	Humorous	----	Tone tends to be neutral, evoking Grok to ridicule, satirise, mock or joke about a politician or party.	@Grok is this a parody? @Grok tell me about [Party leader X] as if you were comedian [Y]
	Innuendo		Tone is always undetermined. The interaction may aim to consolidate a personal or collective view regarding a politician or political proposal, but any intent remains ambiguous as it is primarily an implicit suggestion rather than explicit evaluation	@grok can you explain to me why after all the revelations surrounding [Party Leader X], he can still get Y% of the votes? @grok [Party Leader X] writes for the [Y] newspaper. Which group owns this newspaper? Are there any news reports about the administrators' paychecks?

Source: Own elaboration

The *Tone* dimension includes categories for neutral, positive and negative tone, as well as a fourth category for undetermined, introduced to capture a recurring pattern observed during the coding process. While the boundaries distinguishing positive or negative tone were relatively consensual among coders, a substantial number of posts presented a more nuanced framing. These interactions were frequently articulated through seemingly neutral questions. Yet, these carried implicit cues that complicated a straightforward tone classification. This observation motivated the creation of an undetermined category.

These undetermined interactions tend to be ambiguous on whether the user is aiming to form an opinion or to attack a political opponent. Such posts often contain innuendo or indirect attempts to condition the AI to offer certain responses, frequently targeting a specific politician or party, but with the absence of explicit evaluative markers. These features differentiate them from posts classified under the negative tone category, which are defined by explicit critical language.

This pattern was particularly salient in the context of highly polarising political topics, such as the debate around rents ceilings. In several cases, repeated prompts directed at Grok appeared to seek confirmation of a pre-existing viewpoint, while maintaining the surface form of neutral inquiry. Rather than treating these undetermined interactions as analytically residual, we conceptualised tonal indeterminacy as informative in itself, as it appears to be signaling an underlying communicative strategy. For this reason, posts classified as undetermined in tone were analysed at the level of Manifest Intent and systematically identified as innuendo. In this framework, innuendo is not treated as a tonal category, but as a discursive mode that accounts for how political meaning is conveyed when explicit evaluative cues are absent.

The definition of the humorous posts was also widely discussed among coders. Despite the conceptual breadth of humour in political discourse, consensus was reached that this category should be restricted to posts that are clearly not serious (see examples in table 1).

The *Manifest Intent* dimension includes six parameters, related to: a) the building of knowledge aimed at forming a political position, b) ideological stereotyping, c) moralising through references of personal / political misconduct, d) issues relating to electoral metrics and performance and e) humour. The category “innuendo” has a direct correspondence to the undermined tone.

The use of content analysis procedures (Krippendorff, 2019) allowed us to create a systematic and replicable coding procedure focused on *Manifest Intent* (what the user asks) but which also includes the latent meaning (why / how it is asked). Due to the complex nature of the content at hand and to the recency of AI integration on social media, of the Grok AI features, this approach helped sustain the definition of precise and mutually exclusive categories.

This exercise is particularly challenging in heavily politicised networked digital environments due to the heavy reliance on discursive strategies such as irony, satire or the use of an intentionally manipulative tone (Marwick & Lewis, 2017). This limitation is particularly salient when considering the uses of Grok, which is known for its more irreverent tone and for answering questions that other chatbots refuse to answer or answer in a more politically correct manner (X Corp, n.d).

To facilitate the qualitative appreciation and evaluation of the tweets, in the *Tone* dimension, several categorial variations were defined within the Negative tone category, as all three coders referenced that during initial coding negative toned posts were more complex to categorise. Although not being part of the ICR procedures for the purpose of this exercise, the

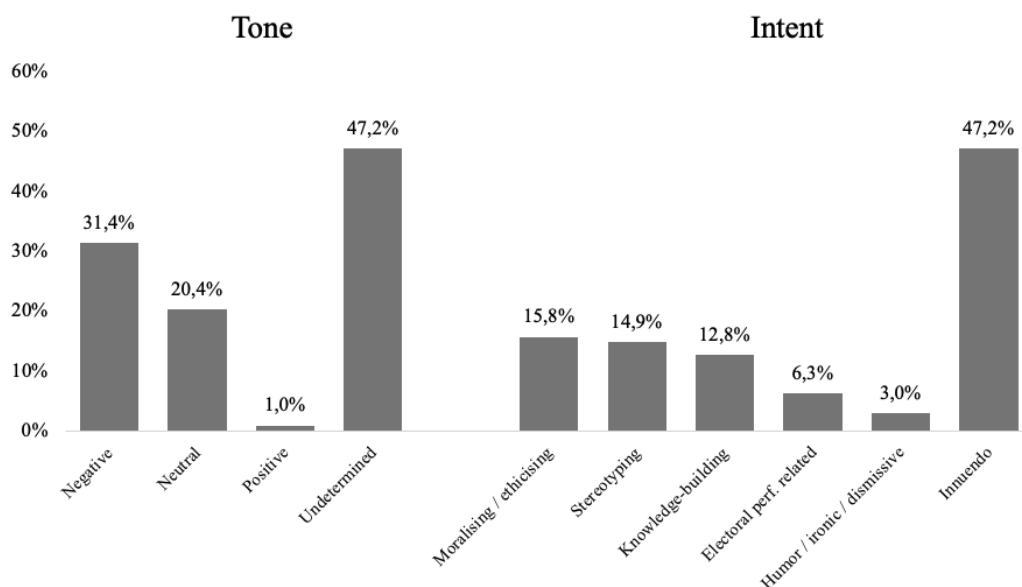
role of the categorical variations was to facilitate the coding of *Tone* dimensions, and to provide depth in terms of the qualitative analysis of results.

After the initial coding of 20% of the database, Krippendorff’s Alpha was employed to assess the inter-rater reliability (Krippendorff, 2019) using the web-based statistical package K-Alpha Calculator (Marzi et al., 2024). The resulting Krippendorff’s Alpha for the dimension of *Tone* is 0.70 (95% confidence interval: 0.61, 0.81), and the coefficient for the dimension of *Manifest intent* is 0.82 (95% confidence interval: 0.66, 0.94).

The threshold for a satisfactory level of this coefficient is 0.80 (Krippendorff, 2019) and even though this threshold was met for the dimension of *Manifest intent* all three coders coded the whole database in both dimensions (525 tweets). Discrepancies were consensualised by a rule of majority of two coders to one. Due to an initial lack of consensus between coders, 36 posts were recoded through a final collective discussion of the codebook.

4. Results

Figure 1. Distribution of replies to politicians or political parties using X Grok by Tone and Intent, Portugal, March 7th 2025 to May 18th 2025 (%)



N=525. Source: Own elaboration.

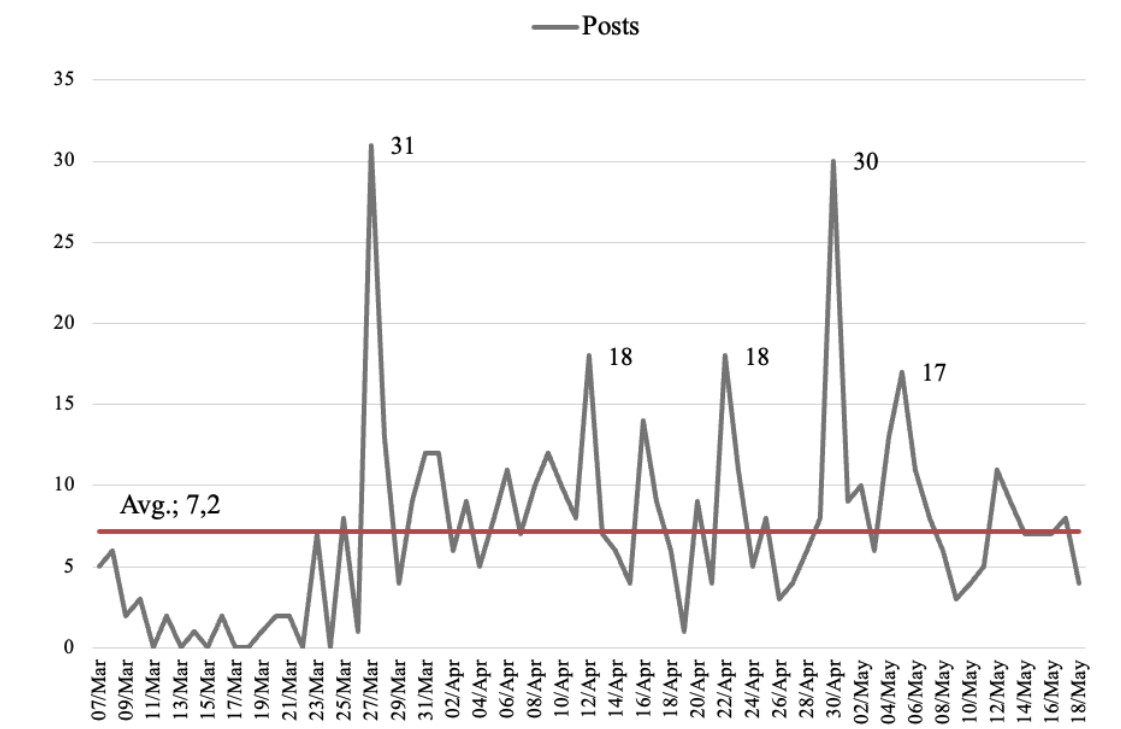
Over a period of 73 days, a total of 525 replies to posts by political parties and leaders were identified in which users explicitly invoked Grok within the discussion. Figure 1 presents the overall distribution of replies according to tone and manifest intent. Regarding tone, 31.4% of posts were coded as negative (n=165), 20.4% as neutral (n=107), and 1.0% as positive (n=5). A substantial proportion of posts (47.2%, n=248) were classified as undetermined in tone, reflecting cases where tone could not be clearly identified based on the available textual content.

In terms of manifest intent, the largest category corresponds to Innuendo (47.2%, n=248) followed by moralising (15.8%, n=83), stereotyping (14.9%, n=78), and knowledge-building (12.8%, n=67). Electoral performance related posts account for 6.3% (n=33), while humorous content represents 3.0% (n=16).

These distributions indicate that, alongside clearly identifiable categories of intent such as moralising or knowledge-building, a large share of interactions do not conform to explicit or unambiguous communicative purposes, despite the clear attempt at a suggestion aimed at the

target using Grok. The coexistence of informational requests and evaluative or expressive uses is consistent with prior research on hybrid forms of political communication in social media environments (Chadwick, 2017; Vaccari & Valeriani, 2021).

Figure 2. Number of X Grok posts in reply to politicians or political parties by date, Portugal, March 7th, 2025, to May 18th, 2025 and description of activity in peak moments



Date of peak	Description
March 27 th (n=31)	High activity around Mariana Mortágua (BE, n=9) and Rui Tavares (Livre, n=8), negative in tone, with stereotyping and moralization intent in remarks.
April 12 th (n=18)	High activity around Mariana Mortágua (BE, n=11) with negative tone and moralizing intent – Posts about candidate’s lifestyle which does not match the far-left egalitarian ideology she defends.
April 22 nd (n=18)	High activity around the PCP – Communist Party (n=10) with negative tone and moralizing intent – Post about party defends humanist values and attacking far-right ideology while refusing to acknowledge the consequences of the communist regimes.
April 30 th (n=30)	High activity around Pedro Nuno Santos and PS (n=13) negative in tone with moralizing intent – Posts about candidate being incompetent / unethical, his proposals do not match his past actions as a minister.
May 5 th (n=17)	High activity around several candidates and parties, namely candidates Pedro Nuno Santos, André Ventura, Mariana Mortágua and parties PCP, BE and PS. Post typology not outside the general trends identified in Figure 1, just a peak in quantity of posts.

N=525. Source: Own elaboration.

Figure 2 displays the temporal distribution of replies invoking Grok between March 7th and May 18th, 2025. Over this period, the average number of posts per day is 7,2, with a median of 7 (minimum=1; maximum= 31; $\sigma= 5.8$), indicating a variable but sustained level of activity.

In the initial phase, between March 7th and March 26th, the number of posts is comparatively lower, with an average of 2.1 posts per day. From late March onwards, activity increases, with several peaks observed on March 27th, April 12th, April 22nd, April 30th, and May 5th. During the official campaign period (May 4th–May 18th), the average number of daily posts is approximately 8.

The table below Figure 2 identifies five peak activity dates in Grok-related replies, detailing volume and thematic focus. Peaks are associated with specific political figures and parties, often involving negative tone and moralising or stereotypical intent. Activity intensifies around prominent actors, with variations reflecting episodic engagement rather than consistent thematic shifts across the observation period.

Table 2. Tone and Manifest intent in using X Grok to reply to politicians or political parties, Portugal, March 7th, 2025, to May 18th, 2025 (%)

Tone							
	Negative	Neutral	Positive	Undetermined	Total		
AndreCVentura (n=164)	26.8%	42.7%	0.6%	29.9%	100%		
MRMortagua (n=88)	27.3%	6.8%	0.0%	65.9%	100%		
PNSpedronuno (n=41)	29.3%	2.4%	0.0%	68.3%	100%		
Ruitavares (n=17)	23.5%	23,5%	0.0%	52.9%	100%		
Ruirochaliberal (n=13)	15.4%	0.0%	7.7%	76.9%	100%		
LMontenegro_PT (n=4)	0.0%	0.0%	0.0%	100.0%	100%		
InesSousaReal (n=1)	0.0%	0.0%	0.0%	100.0%	100%		
PartidoCHEGA (n=40)	35.0%	40,0%	0,0%	25,0%	100%		
BlocoDeEsquerda (n=54)	33.3%	1,9%	0,0%	64,8%	100%		
Psocialista (n=29)	37.9%	3.4%	3.4%	55.2%	100%		
LIVREpt (n=12)	8.3%	8.3%	0.0%	83.3%	100%		
LiberalPT (n=21)	33.3%	19.0%	0.0%	47.6%	100%		
Pcp_pt (n=41)	68.3%	7.3%	4.9%	19.5%	100%		
Manifest Intent							
	Moralising or ethicising	Stereotyping	Knowledge- building	Electoral performance related	Humorous	NA (Undetermined Tone)	Total
AndreCVentura (n=164)	17.1%	8.5%	24.4%	15.9%	4.3%	29.9%	100%
MRMortagua (n=88)	19.3%	8.0%	4.5%	1.1%	1.1%	65.9%	100%
PNSpedronuno (n=41)	22.0%	7.3%	0.0%	0.0%	2.4%	68.3%	100%
Ruitavares (n=17)	17.6%	0.0%	17.6%	0.0%	11.8%	52.9%	100%
Ruirochaliberal (n=13)	15.4%	0.0%	0.0%	0.0%	7.7%	76.9%	100%
LMontenegro_PT (n=4)	0.0%	0.0%	0.0%	0.0%	0.0%	100%	100%
InesSousaReal (n=1)	0.0%	0.0%	0.0%	0.0%	0.0%	100%	100%
PartidoCHEGA (n=40)	17.5%	12.5%	27.5%	10.0%	7.5%	25.0%	100%
BlocoDeEsquerda (n=54)	14.8%	18.5%	1.9%	0.0%	0.0%	64.8%	100%
Psocialista (n=29)	24.1%	17.2%	3.4%	0.0%	0.0%	55.2%	100%
LIVREpt (n=12)	0.0%	8.3%	8.3%	0.0%	0.0%	83.3%	100%
Pcp_pt (n=41)	4.9%	65.9%	4.9%	4.9%	0.0%	19.5%	100%
LiberalPT (n=21)	0.0%	28.6%	19.0%	0.0%	4.8%	47.6%	100%

Note: heatmaps relate to the numerical interval comprising posts that are either Negative, Neutral or Positive, excluding posts categorised as Undetermined.

n=525. Source: Own elaboration.

In the upper half, Table 2 presents the distribution of tone across replies directed at individual political leaders and their party accounts. Across all actors, the positive tone remains marginal to non-existent. Negative and undetermined tones dominate the distribution, although their relative proportions vary.

Among political leaders, replies to André Ventura (n=164) show a distribution of 26.8% negative, 42.7% neutral, and 29.9% undetermined. For Mariana Mortágua (n=88) and Pedro Nuno Santos (n=41), undetermined tone constitutes most cases (65.9% and 68.3%, respectively). Other leaders display similar patterns, although with smaller sample sizes.

At the party level, differences in tone distribution are also observable. PCP (n=41) presents the highest proportion of negative tone (68.3%), while Chega (n=40) shows a higher proportion of neutral tone (40.0%) compared to other parties. Bloco de Esquerda (n=54) and Partido Socialista (n=29) exhibit a predominance of undetermined tone (64.8% and 55.2%, respectively).

On the bottom half of Table 2, we find the details the distribution of Manifest intent across actors and parties. Knowledge-building and electoral performance-related posts are more frequent in replies directed at André Ventura (24.4% and 15.9%, respectively) and Chega (27.5% and 10.0%). Moralising intent appears more frequently in replies directed at Partido Socialista (24.1%), Pedro Nuno Santos (22.0%), and Mariana Mortágua (19.3%). Stereotyping is particularly prominent in replies to PCP (65.9%) and, to a lesser extent, Iniciativa Liberal (28.6%).

These distributions illustrate variation in the types of interactions associated with different political actors, reflecting diverse uses of Grok across contexts of engagement. In addition to the categorical analysis of tone and manifest intent, the data reveals recurring configurations in how users mobilise Grok within political interactions. These configurations emerge from the combination of tone and intent dimensions and are observable across multiple cases in the dataset. Four main use models were identified.

Table 3. Main usage models of @Grok in replies to politicians or political parties in Portugal on X ahead of the Portuguese legislative elections of May 2025

Use model	Tone and manifest intent
Opinion validation	Moralising or stereotyping intent with negative tone and innuendo with undetermined tone
Argumentation by proxy	Moralising or stereotyping intent with negative tone
Constructive autonomisation	Neutral tone and knowledge-building or electoral performance-related intent.
Discursive and memetic framing	Better identified under humours intent with apparent neutral tone but tone and intent may vary

Source: Own elaboration.

First, *opinion validation* describes instances where users formulate questions or comments that align with moralising or stereotyping intent, often in conjunction with negative tone, or reveal a more nuanced approach with innuendo as a discursive marker and an undetermined tone. These interactions involve the use of Grok in ways that correspond to the reinforcement of pre-existing positions or perspectives by evoking an anticipated Grok response on salient political issues.

Second, *argumentation by proxy* refers to cases in which users construct prompts that condition the AI into generating evaluative or critical statements itself. These instances are typically associated with negative tone and stereotyping intent and involve the externalisation of argumentative content through the AI interface. It appears to be used to avoid accountability of controversial or extreme statements while amplifying polarised discourse.

Third, *constructive autonomisation* refers to instances where users invoke Grok as an informational resource, typically associated with neutral tone and knowledge-building or electoral performance-related intent. These posts involve requests for clarification, verification, or contextualisation of political information, therefore involving the use of AI for fact-checking practices.

Finally, *discursive and memetic framing* encompasses interactions in which Grok is used as part of broader communicative practices characteristic of the platform, including humour, irony, and symbolic expression. In these cases, tone and intent may vary, but the use of Grok aligns with established patterns of memetic and performative engagement on X.

These use-models are not mutually exclusive and may overlap within individual posts. Rather than representing fixed categories, they capture recurrent patterns in the ways users integrate AI into political discourse. Additionally, the articulation between models of usage, tone, and manifest intent, as shown in Table 3, is based on our observation of posts in the context of the 2025 Portuguese Legislative Elections. However, these models of usage could potentially be applied across different contexts, with varying tones and intents.

5. Discussion

During the taxonomisation of posts and analysis of results several relevant aspects arose, namely regarding the relationship between tone, manifest intent, and the broader dynamics of political communication on X. Despite the novelty of AI-powered social media engagement between users for political debate, the results can be read within the broader theoretical scope of social media studies and AI-mediated interaction.

A first notable aspect concerns the overall distribution of tone and intent. The predominance of negative and undetermined tone, combined with the relatively limited presence of positive or supportive interactions, indicates that Grok is primarily mobilised within contentious or ambiguous communicative contexts. This pattern is consistent with research highlighting the prominence of adversarial, polarised, and emotionally charged discourse on social media platforms, particularly in political settings (Tufekci, 2015, Tucker et al., 2018). Simultaneously, the coexistence of moralising, and stereotyping intent on one side, and knowledge-building intent on another, suggests that Grok is integrated into a complex communicative environment where informational, evaluative, and expressive functions coexist (Chadwick, 2017, Vaccari & Valeriani, 2021).

The high proportion of cases classified as “undetermined” in tone or “innuendo” in manifest intent is particularly significant. While this category reflects the limits of inferring tone and intent solely from textual content between different coders, it also points to a broader characteristic of AI-mediated political discourse. Many interactions mentioning Grok are formulated in ways that rely on ambiguous, indirect or implied meaning, making it difficult to attribute a clear communicative purpose. This observation aligns with previous studies emphasizing the role of irony, innuendo, and performative ambiguity in networked political communication (Highfield, 2016, Marwick & Lewis, 2017, Philips & Milner, 2017).

In this sense, the prevalence of undetermined cases should not be understood only as a limitation of the coding process, but also as indicative of a communicative style in which meaning is often negotiated implicitly rather than explicitly. Two factors are worth considering in evaluating the meaning of the large proportion of undetermined tone / innuendo-filled intent posts. The first is the fact that at the time of the post collection process the @Grok feature was very recent and its nature as an affordance was still being apprehended. The second relates to the possibility that in the case of X ambiguity, implicit meaning and indirectness are a feature

of non-regulated space for political debate lying outside the normative frameworks of any other political mediated arenas which tend to be subject to more institutionalised and normative frameworks of mediation. This relative absence of regulation not only enables, but actively invites, practices of exploratory and strategically ambiguous engagement, consistent with what X describes as the "rebellious streak" of Grok (X Corp, n.d.), that position the platform against mainstream norms of moderation and accountability. Within this context, users may adopt a rhetoric of neutrality in form, while, in practice seeking to to validate pre-existing opinions through Grok's expected response, particularly on prominent political issues.

The distribution of interactions across political actors further illustrates how the use of Grok varies depending on context and target. Differences observed in tone and intent across parties and leaders suggest that AI-mediated engagement is not uniform but shaped by patterns of visibility, issue prominence, and political positioning. For instance, the presence of knowledge-building and electoral performance-related posts in replies directed at certain actors coexists with moralising and stereotypical content directed at others.

In this regard it is notable that André Ventura has the lowest percentage of undetermined replies among political leaders, which suggests that users tend to engage with his posts in a less ambiguous manner, probably due to its populist discourse that tends to divide people. On the other hand, and even though negativity dominates interactions, especially toward the Communist party, ambiguity remains a central feature in Grok-powered engagement with many political actors.

Such variation is consistent with research showing that user engagement on platforms like X is influenced by both the communicative strategies of political actors and the interpretative frameworks of audiences (Waisbord & Amado, 2017; Marwick & Lewis, 2017). In this light it is worth pointing out that political parties are themselves experimenting with AI, while not being held responsible for its usage, and therefore have little incentive to reject the technology (Jungherr et al., 2026).

Beyond individual distributions, the existence of recurring patterns of Grok usage provides additional insight into how AI is incorporated into political interactions. The four use models identified in the Results (*opinion validation*, *argumentation by proxy*, *constructive autonomisation* and *discursive and memetic framing*) capture distinct but overlapping ways in which users mobilise the AI within the platform's communicational environment. The identification of these patterns reflects the hybrid nature of AI-mediated communication in a social platform environment, combining informational, expressive, and performative elements within the same interactional space (Papacharissi, 2014). In this sense, the addition of an AI powered third-party to the platform generates a new framework of affordances that mutates the functional and interactional possibilities in the environment (Recuero, 2025).

Answering RQ1., ('How was Grok used to interact with posts made by political parties and leaders on X ahead of the 2025 legislative elections in Portugal?') We find that users are very active in the mobilisation of Grok as a tool and do it in four main ways.

Through *constructive autonomisation* they actively gather or verify information, most often in connection with knowledge-building or electoral performance-related queries. This matches the growing role of users as active participants in information verification processes, particularly in contexts where traditional fact-checking infrastructures are complemented by digital tools (Graves, 2016, Juneja & Mitra, 2022). Evidence also points to a markedly performative aspect in the use of Grok, better understood using *discursive and memetic framings*. Beyond the informational use-case the AI is part of a broader communicative repertoire involving humour, irony, and shared cultural references. This finding aligns with research on

meme culture and the role of platforms in shaping collective forms of expression and identity (Galip, 2024), as well as with the idea of "affective publics" in digitally mediated political communication (Papacharissi, 2014) and in platforms where discussions are meant to be seen, and beyond the user's sphere (Murphy, 2018, Bruns, 2021).

Observing the results and considering the overwhelming amount of negative and neutral posts, one core aspect of the answer to RQ1. must address the patterns described as *opinion validation* and *argumentation by proxy*, linked to evaluative and critical stances. Here, evidence points to the formulation of prompts aiming to condition the AI into a desirable argumentative outcome. While this may be seen by users as an argumentative resource, it also highlights the potential for AI systems to reinforce existing perspectives or to be integrated into polarised communicative environments (Sharma et al., 2023; Chen et al., 2025).

Chatbots, when utilized under appropriate circumstances, can become useful resources within the repertoire of tools used for fact-checking purposes (Kuznetsova et al., 2025) improving user knowledge and therefore their status (Papacharissi, 2014, Benkler et al., 2018, Farkas & Schou, 2023). Grok is a powerful assistant in empowering users while also making them more vulnerable to accepting the arguments of others as valid and to overturning their own bias (Martínez-Costa et al., 2022, Sharma, 2023, RRV, 2024, Chen, 2025). This effect may be even pronounced as Grok's good benchmarks in certain technical tasks and topics (Jedrzejczak et al., 2025) is often conflated with an outstanding ability of the AI to verify facts.

Taken together, these findings suggest that the role of Grok in political discourse extends beyond its potential function as an informational or fact-checking tool. Instead, it operates within a broader context of communication in which informational, discursive, and performative elements are interconnected in a digital platform structure that does not simply allow the exchange of information but also shapes the configurations of political interaction (van Dijck et al., 2018, Recuero, 2025).

6. Conclusion

This study focuses on how users mobilise AI, specifically the Ask Grok feature, within online political discussions on X during the 2025 Portuguese Legislative election period. Grok stands out within the contemporary AI landscape not only because of its technical capabilities, but also due to its integration within X and its positioning as a visible participant in ongoing conversations.

Unlike standalone chatbot interfaces, Grok operates within the flow of public day-to-day interaction, enabling users to incorporate AI directly into exchanges with political actors and with each other. This feature, combined with the platform's established culture of direct engagement, contributes to a distinctive communicative environment in which AI becomes part of the performative and interactive dynamics of political discourse, as well as a key resource in helping gather knowledge to help manifest sustained political opinions, aligned with X goals.

Although X has relatively low penetration in Portugal, it remains a highly active and consequential space for political communication, particularly among politically engaged users. The electoral context further intensifies this dynamic, providing a setting in which interactions are more frequent, visible, and contested.

Within this environment, the findings show that Grok is mobilised in four main ways, namely (i) *opinion validation*, (ii) *argumentation by proxy*, (iii) *constructive autonomisation* and (iv) *discursive and memetic framing*. This mode of use was based on limited evidence of explicitly

positive or supportive uses, but a wide range of intentions spanning knowledge-building, moral evaluation, stereotyping, humour, and electoral performance queries.

A central finding of the study concerns the high proportion of cases classified as “undetermined” (tone) / “innuendo” (manifest intent). While this initially presents a methodological challenge, it also constitutes a substantive finding into the nature of AI-mediated political discourse. Many interactions invoking Grok are characterised by ambiguity, indirectness, and the use of irony or implicit meaning, which complicate straightforward interpretation. Rather than being an artifact of coding limitations alone, this ambiguity might reflect the prevalence of performative and layered communication practices on X, where meaning is often implied rather than explicitly stated.

At the same time, the findings should be considered in light of some limitations. The dataset captures a specific platform, a defined period, and a particular national context, which constrains the generalisability of the results. Moreover, the reliance on textual content limits the ability to fully interpret user intent, especially in cases characterised by ambiguity or indirect expression. These limitations reinforce the exploratory nature of the study and highlight the need for a proper contextual interpretation when analysing AI-mediated interactions in political contexts.

In this sense, the study contributes to the emerging literature on AI-mediated political communication by shifting attention from the accuracy or effectiveness of AI systems to their role within broader communicative practices. The findings suggest that AI is not only used as a tool for information retrieval or verification, but also as part of a wider repertoire of expressive, symbolic, and performative actions shaped by platform affordances and user cultures. This perspective helps situate AI within the evolving dynamics of digital political communication, where meaning is co-constructed through interaction rather than simply transmitted.

Future research can build on this work by expanding the scope of analysis across platforms, contexts, and types of AI systems, as well as by employing more advanced quantitative and mixed-method approaches. Such efforts may help to further disentangle the relationships between tone, intent, and actor-specific dynamics, while also contributing to a more comprehensive understanding of how AI is integrated into contemporary political discourse.

Declaration of interest

The authors declare that they have no personal or commercial interests or relationships that could pose a conflict of interest with respect to the submitted article. This research is original, has been submitted to *Communication & Society* only and follows strict ethical principles and guidelines of research.

Acknowledgments

This publication has been co-funded by the European Commission under the Call DIGITAL2023 DEPLOY-04, European Digital Media observatory (EDMO) – National and multinational hubs (reference: IBERIFIER Plus – 101158511). The authors thank the reviewers for their suggestions that contributed to the quality of this article. We thank the guest editors of this special issue and editorial team of *Communication & Society* for their tireless and outstanding efforts in facilitating the peer-review and publication processes. We also thank our colleague Nuno Palma for his help in extracting the data that supports this article.

References

- oDay3xploit (2025). ZenScraper. GitHub. <https://github.com/oDay3xploit/ZenScraper> (Accessed on September 2025).
- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–236. <https://doi.org/10.1257/jep.31.2.211>
- Barrie, C. (2022). Did the Musk takeover boost contentious actors on Twitter? arXiv preprint arXiv:2212.10646. <https://doi.org/10.48550/arXiv.2212.10646>
- Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press. <https://doi.org/10.1093/oso/9780190923624.001.0001>
- Bennett, W. L., & Segerberg, A. (2012). The logic of connective action: Digital media and the personalization of contentious politics. *Information, Communication & Society*, 15(5), 739–768. <https://doi.org/10.1080/1369118X.2012.670661>
- Benton, B., Choi, J. A., Luo, Y., & Green, K. (2022). Hate speech spikes on Twitter after Elon Musk acquires the platform. School of Communication and Media, Montclair State University. <https://digitalcommons.montclair.edu/scom-facpubs/33>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Bruns, A. (2012). How long is a tweet? Mapping dynamic conversation networks on Twitter using Gawk and Gephi. *Information, Communication & Society*, 15(9), 1323–1351. <https://doi.org/10.1080/1369118X.2011.635214>
- Cardoso, C., Moreno, J., Narciso, I., Couraceiro, P., & Santos, J. (2025a). *Legislativas 2025 – Informação e desinformação nas redes sociais*. OberCom – Observatório da Comunicação. <https://doi.org/10.5281/zenodo.15781756>
- Cardoso, G., Paisana, M., & Pinto-Martinho, A. (2025b). *Digital News Report Portugal 2025*. OberCom – Observatório da Comunicação. <https://doi.org/10.5281/zenodo.15599071>
- Chadwick, A. (2017). *The hybrid media system: Politics and power* (2nd ed.). Oxford University Press. <https://doi.org/10.1093/oso/9780190696726.001.0001>
- Charmaz, K. (2006). *Constructing Grounded Theory*. London: Sage.
- Chen, S., Gao, M., Sasse, K., Hartvigsen, T., Anthony, B., Fan, L., Aerts, H., Gallifant, J., & Bitterman, D. S. (2025). When Helpfulness Backfires: LLMs and the Risk of Misinformation Due to Sycophantic Behavior. *Research square*, rs.3.rs-6206365. <https://doi.org/10.21203/rs.3.rs-6206365/v1>
- Chehoudi, R. (2025). Artificial intelligence and democracy: pathway to progress or decline? *Journal of Information Technology & Politics*, 1–16. <https://doi.org/10.1080/19331681.2025.2473994>
- Christopher, N., & Pepe, V. (2025, July 11). As millions adopt Grok to fact-check, misinformation abounds. *Al Jazeera*. <https://www.aljazeera.com/economy/2025/7/11/as-millions-adopt-grok-to-fact-check-misinformation-abounds>
- Chuai, Y., Tian, H., Pröllochs, N., & Lenzini, G. (2024). Did the roll-out of community notes reduce engagement with misinformation on X/Twitter? *Proceedings of the ACM on human-computer interaction*, 8(CSCW2), 1–52. <https://doi.org/10.1145/3686967>
- Drolsbach, C. P., & Pröllochs, N. (2023). Diffusion of community fact-checked misinformation on twitter. *Proceedings of the ACM on Human-Computer Interaction*, 7(CSCW2), 1–22. <https://doi.org/10.48550/arXiv.2205.13673>
- European Commission: Directorate-General for Communications Networks, Content and Technology (2025). *Code of conduct on disinformation: as amended in October 2024*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2759/5029213>
- Farkas, J., & Schou, J. (2023). *Post-Truth, Fake News and Democracy: Mapping the Politics of Falsehood* (2nd ed.). Routledge. <https://doi.org/10.4324/9781003434870>

- Ferro-Santos, S., Cardoso, G., & Santos, S. (2024). What do Portuguese MPs use Twitter for? A case study on political communication in a country with a low Twitter adoption rate. *Cuadernos.info*, (58), 184–207. <https://doi.org/10.7764/cdi.58.72109>
- Fortune. (2025, July 10). More than 90 % of X's Community Notes are never published and 'stuck in limbo,' study finds. *Fortune*. <https://fortune.com/2025/07/10/more-than-90-x-community-notes-are-never-published-stuck-limbo-study/>
- Galip, I. (2024). Methodological and epistemological challenges in meme research and meme studies. *Internet Histories*, 8(4), 312–330. <https://doi.org/10.1080/24701475.2024.2359846>
- Glaser, B. G., & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine.
- Gonçalves, V. (2023). Portuguese Party Leaders on Twitter: Interactions and Media Hybridisation Strategies. *Comunicação e sociedade*, (44), 1–28. [https://doi.org/10.17231/comsoc.44\(2023\).4828](https://doi.org/10.17231/comsoc.44(2023).4828)
- Graves, L. (2016). *Deciding what's true: The rise of political fact-checking in American journalism*. Columbia University Press.
- Highfield, T. (2016). *Social Media and Everyday Politics*. Polity.
- Isaac, M. (2025, July 8). Elon Musk's Grok chatbot shares antisemitic posts on X. *The New York Times*. <https://www.nytimes.com/2025/07/08/technology/grok-antisemitism-ai-x.html>
- Islam, M. B. E., Haseeb, M., Batool, H., Ahtasham, N., & Muhammad, Z. (2024). AI Threats to Politics, Elections, and Democracy: A Blockchain-Based Deepfake Authenticity Verification Framework. *Blockchains*, 2(4), 458–481. <https://doi.org/10.3390/blockchains2040020>
- Jedrzejszak, W. W., Szkielkowska, A., Raj-Koziak, D., Włodarczyk, E., Skarżyński, H., & Kochanek, K. (2025). How easily can AI chatbots spread misinformation in audiology and otolaryngology? *medRxiv*, 2025-04. <https://doi.org/10.1101/2025.04.24.25326281>
- Jude, N., & Matamoros-Fernández, A. (2025, February 3). Community Notes and its narrow understanding of disinformation. *TechPolicy.Press*. <https://www.techpolicy.press/community-notes-and-its-narrow-understanding-of-disinformation/>
- Juneja, P., & Mitra, T. (2022). Human and technological infrastructures of fact-checking. *arXiv*. <https://doi.org/10.48550/arXiv.2205.10894>
- Jungherr, A., Rauchfleisch, A., & Wuttke, A. (2026). Artificial Intelligence in Election Campaigns: Perceptions, Penalties, and Implications. *Political Communication*, 1–22. <https://doi.org/10.1080/10584609.2025.2611913>
- Krippendorff, K. (2019). *Content Analysis: An Introduction to Its Methodology* (4th Ed.). SAGE Publications. <https://doi.org/10.4135/9781071878781>
- Kuznetsova, E., Makhortykh, M., Vziatysheva, V., Stolze, M., Baghumyan, A., & Urman, A. (2025). In generative AI we trust: can chatbots effectively verify political information? *Journal of Computational Social Science*, 8(15). <https://doi.org/10.1007/s42001-024-00338-8>
- Lewandowsky, S., Ecker, U. K., Cook, J., Van Der Linden, S., Roozenbeek, J., & Oreskes, N. (2023). Misinformation and the epistemic integrity of democracy. *Current opinion in psychology*, 54, Article 101711. <https://doi.org/10.1016/j.copsyc.2023.101711>
- Li, J., & Chang, X. (2022). Combating Misinformation by Sharing the Truth: a Study on the Spread of Fact-Checks on Social Media. *Information systems frontiers: A journal of research and innovation*, 1–15. Advance online publication. <https://doi.org/10.1007/s10796-022-10296-z>
- López-García, G. (2016). Nuevos y 'viejos' liderazgos: La campaña de las elecciones generales españolas de 2015 en Twitter. *Comunicación y Sociedad*, 29(3), 149–167. <https://doi.org/10.15581/003.29.3.sp.149-167>
- Martínez-Costa, M. P., López-Pan, F., Buslón, N., & Salaverría, R. (2022). Nobody-fools-me perception: Influence of Age and Education on Overconfidence About Spotting Disinformation. *Journalism Practice*, 17(10), 2084–2102. <https://doi.org/10.1080/17512786.2022.2135128>

- Marwick, A. and Lewis, R. (2017) Media Manipulation and Disinformation Online. Data & Society Research Institute. <https://datasociety.net/library/media-manipulation-and-disinfo-online/>
- Marzi, G., Balzano, M., & Marchiori, D. (2024). K-Alpha Calculator—Krippendorff's Alpha Calculator: A User-Friendly Tool for Computing Krippendorff's Alpha Inter-Rater Reliability Coefficient. *MethodsX*, 12, 102545. <https://doi.org/10.1016/j.mex.2023.102545>
- Mehta, I. (2025, March 7th). X now lets you query Grok by mentioning it in replies. *TechCrunch*. <https://techcrunch.com/2025/03/07/x-now-lets-you-query-grok-by-mentioning-it-in-replies/>
- Murphy, D. (2018). *Twitter*. Cambridge, Polity.
- Newman, N., Ross Arguedas, A., Robertson, C. T., Nielsen, R. K., & Fletcher, R. (2025). Digital news report 2025. Reuters Institute for the Study of Journalism. <https://dx.doi.org/10.60625/risj-8qqf-jt36>
- Nisbet, E. C., Mortenson, C., & Li, Q. (2021). The presumed influence of election misinformation on others reduces our own satisfaction with democracy. *Harvard Kennedy School (HKS) Misinformation Review*, 1(7). <https://doi.org/10.37016/mr-2020-59>
- OECD (2024), *Facts not Fakes: Tackling Disinformation, Strengthening Information Integrity*, OECD Publishing, Paris, <https://doi.org/10.1787/d909ff7a-en>
- Ognyanova, K., Lazer, D., Robertson, R. E., & Wilson, C. (2020). Misinformation in action: Fake news exposure is linked to lower trust in media, higher trust in government when your side is in power. *Harvard Kennedy School (HKS) Misinformation Review*. <https://doi.org/10.37016/mr-2020-024>
- Paisana, M., Vasconcelos, A., Couraceiro, P., Caldeira Pais, P. M. P., Cardoso, G., Baldi, V., Barriuso Varela, O., Álvarez Corona, E., Espiritusantu, O., Dinant, Ì., Magallon-Rosa, R., Fernández-Castrillo, C., Herrero Curiel, E., Molina Cañabate, J. P., Camacho, D., D'Antonio-Maceiras, S., Huertas-Tato, J., Girón, A., Martín, A., Crespo, M., Zafra, J., Bonilla, L., & Sales-Giménez, M. (2025). *An Iberian Outlook: Report on Social Media Disinformation in Spanish, Portuguese and EU elections and detection tools*. SmartVote. <https://doi.org/10.5281/zenodo.15305356>
- Papacharissi, Z. (2014). *Affective Publics: Sentiment, Technology, and Politics*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/978019999736.001.0001>
- Philips, W., & Milner, R.M. (2017). *The Ambivalent Internet: Mischief, Oddity, and Antagonism Online*. Wiley.
- Qian, Y., Siau, K., & Nah, F.F. (2024). Societal impacts of artificial intelligence: Ethical, legal, and governance issues. *Societal Impacts*, 3. <https://doi.org/10.1016/j.socimp.2024.100040>
- Recuero, R. (2025). A systemic framework for disinformation on social media platforms. *Platforms & Society*, 2. <https://doi.org/10.1177/29768624251367199>
- Roa, E.G. (2024). Impact of artificial intelligence. Opportunity or threat. New markets, new business. Are we facing a new industrial revolution? [Bachelor's thesis, Universitat Politècnica de Catalunya BarcelonaTech]. <https://hdl.handle.net/2117/418554>
- RRV, A., Tyagi, N., Uddin, M. N., Varshney, N., & Baral, C. (2024). Chaos with keywords: Exposing large language models' sycophantic hallucination to misleading keywords and evaluating defense strategies. *arXiv:2406.03827*. <https://doi.org/10.48550/arXiv.2406.03827>
- Samet, U. (2025). The positive influence of large language models on fact-checking practices: A case study of Grok. *World Journal of Advanced Engineering Technology and Sciences*, 15(03), 1727-1738. <https://doi.org/10.30574/WJAETS.2025.15.3.1123>
- Sharma, M., Tong, M., Korbak, T., Duvenaud, D., Askill, A., Bowman, S. R., Cheng, N., Durmus, E., Hatfield-Dodds, Z., Johnston, S. R., Kravec, S., Maxwell, T., McCandlish, S., Ndousse, K., Rausch, O., Schiefer, N., Yan, D., Zhang, M., & Perez, E. (2023). Towards understanding sycophancy in language models. *arXiv*. <https://doi.org/10.48550/arXiv.2310.13548>

- Suomalainen, K., Nykänen, N., Seeck, H., Kim, Y., & McPherson, E. (2025). Fact-Checking in Journalism: An Epistemological Framework. *Journalism Studies*.
<https://doi.org/10.1080/1461670X.2025.2492729>
- Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D., & Nyhan, B. (2018). Social media, political polarization, and political disinformation: A review of the scientific literature. Hewlett Foundation. <https://dx.doi.org/10.2139/ssrn.3144139>
- Tufekci, Z. (2015). Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency. *Colorado Technology Law Journal*, 13(2), 203–218. Available at: <https://ctlj.colorado.edu/?p=509>
- United Nations (2022). Countering disinformation for the promotion and protection of human rights and fundamental freedoms: report of the Secretary-General. United Nations. <https://digitallibrary.un.org/record/3987886?v=pdf>
- Vaccari, C., & Valeriani, A. (2021). Outside the bubble: Social media and political participation in Western democracies. Oxford University Press.
- van Dijck, J., Poell, T., & de Waal, M. (2018). The platform society: Public values in a connective world. Oxford University Press.
- Waisbord, S., & Amado, A. (2017). Populist communication by digital means: presidential Twitter in Latin America. *Information, communication & society*, 20(9), 1330–1346.
<https://doi.org/10.1080/1369118X.2017.1328521>
- Wojcik, S., Hilgard, S., Judd, N., Mocanu, D., Ragain, S., Hunzaker, M. B., ... & Baxter, J. (2022). Birdwatch: Crowd wisdom and bridging algorithms can inform understanding and reduce the spread of misinformation. arXiv preprint arXiv:2210.15723.
file:///Users/elena/Creative Cloud Files /C&S/Volumen 39/Vol.39 especial/Para autores/ <https://doi.org/10.48550/arXiv.2210.15723>
- X Corp. (2021, January 25). Introducing Birdwatch, a community-based approach to misinformation. Blog – X. Retrieved October 23, 2025, from https://blog.x.com/en_us/topics/product/2021/introducing-birdwatch-a-community-based-approach-to-misinformation
- X Corp. (n.d.). About Grok, your humorous AI assistant on X. Help Center – X. Retrieved October 23, 2025, from <https://help.x.com/en/using-x/about-grok>